VISUALIZING SPATIAL RELATIONSHIPS:

THE EFFECTS OF TWO-DIMENSIONAL, THREE-DIMENSIONAL AND VIRTUAL ENVIRONMENT TRAINING DISPLAYS

Lyn Mowafy

Air Force Research Laboratory, Mesa AZ

Richard Thurman

Air Force Research Laboratory, Mesa AZ

VISUALIZING SPATIAL RELATIONSHIPS:

THE EFFECTS OF TWO-DIMENSIONAL, THREE-DIMENSIONAL AND VIRTUAL ENVIRONMENT TRAINING DISPLAYS

Lyn Mowafy

Richard Thurman

Air Force Research Laboratory, Mesa AZ

For training purposes, there are a variety of means for presenting spatial information to support and augment mental modeling. These include alphanumeric parameter lists, schematic diagrams, plan views and maps, perspective views, and synthesized virtual environments. From a training perspective, we need to know whether a medium contributes significantly to optimizing training resources. It should be implemented in a training program only if it presents useful information with a minimum of distortion, ambiguity, or irrelevant and distracting data. This is particularly true when trainees are learning to construct mental models of spatial relationships from instruments which do not offer a direct or intuitive representation of the necessary information. Under these conditions, the choice of display parameters implemented in a training system will directly effect its utility for aiding in the development of spatial problem solving skills.

The rapid evolution of enabling technologies has provided an array of interactive computer graphic displays, and designers of training tools now enjoy a variety of potentially effective methods for visualizing complex data sets (Fisher & Tazelaar, 1990; Foley, 1987). These include the traditional orthographic, or two-dimensional views, perspective views incorporating three-dimensional pictorial depth cues, and wide angle orthostereoscopic virtual environments. Although the same information can be provided under all of these formats, each embodies unique opportunities and constraints. As a result, the choice of a particular medium for a training program may have a significant impact on the trainee's ability to encode, retain, and recall functionally relevant information. How then do we chose the most appropriate medium?

Task/Display Compatibility

Larkin and Simon (1987) have distinguished between media representations which are informationally equivalent and those which are computationally equivalent. Representations are informationally equivalent if all the information in one is inferable from the other, and vice versa. Computationally equivalent representations, on the other hand, are informationally equivalent and they impose similar attentional and cognitive processing loads when used in problem solving.

Wickens and his colleagues (Liu & Wickens, 1992; Wickens & Andre, 1988; Wickens & Todd, 1990; Wickens, Todd, & Seidler, K. 1989) have assumed a similar theoretical perspective in their approach to display design. They suggest that display media should be tailored to the to-be-learned task according to the proximity compatibility principle. This principle refers to the degree of spatial coincidence, or correspondence, between the task and the format of the display. For example, if the task requires the user to integrate information from multiple sources, the display should present the information in a visually integrated context; perhaps endowing it with object-like properties. On the other hand, tasks requiring the focus of attention on a single dimension of the data would be better served by designing training media to enhance its discriminability (i.e., dissociate that dimension from related information that can cause interference).

This suggests that for tasks requiring the mental representation of spatial relationships in our three-dimensional world, the computational non-equivalence of perspective and orthographic displays would argue in favor of perspective displays. There also is considerable anecdotal evidence to recommend the use of three-dimensional computer graphics in education and training (cf. Bertoline, 1991; Brody, Jacoby & Ellis, 1991; Foley, 1987; Zsombor-Murray, 1990), medical imaging (Russell & Miles, 1989; McConathy, Deirdra & Doyle, 1991) scientific visualization (Gomez, 1989; Farrell & Christidis, 1989), computer-aided design (Greenberg, 1991), and remote teleoperation (Cole & Parker, 1989).

Duality of Displayed Representations

Gibson (1979) has offered a different theoretical approach. His account has focused on describing the nature of the information communicated by the perspective display medium. Gibson has noted that any pictorial medium simultaneously affords two representations. On the one hand, pictures are designed to transport the observer into a "virtual" world. The observer's awareness is of being in another world, or of observing it through a window. To the viewer, that world is not an illusion of reality; nor is it simply a set of unrelated objects. Rather, it is an environment in which objects and actors potentially can move about and interact in a rational, manner. In this case, one's viewpoint into the virtual world can be critical to understanding the spatial relationships depicted. On the other hand, a pictorial display also is an object in the real world. As an object, it's properties consist of a set of spatial relationships among markings on a two-dimensional surface, bounded by the frame of the display. These are not properties of the virtual world; rather they coexist with the virtual world and are accessible directly to the observer at all times.

This duality in pictorial representations creates a dynamic tension which ultimately can impact the utility of a visualization training display - regardless of whether the information conveyed is informationally or computationally equivalent to another display. Indeed, there has been a long-standing controversy regarding the misinterpretation of spatial relationships in perspective images that invokes either the theory of the misapprehended viewpoint into the virtual world of the display or the effect of the two-dimensional frame of reference imposed by the display medium (cf. Goldstein, 1987; Cutting, 1988; Goldstein, 1988). This controversy is relevant to display designers by virtue of the fact that both viewpoint and frame of reference can now be manipulated, if it is determined that these variables are critical to training effectiveness.

The Virtual Worlds of Graphical Display

Ellis (1991) has explored the use of interactive computer graphics to transport the user into Gibson's virtual world beyond the picture frame. He describes the creation of the virtual world as a process of virtualization, "by which a human viewer interprets a patterned sensory impression to be an extended object in an environment other than that in which it physically exists" (Ellis, 1991, p. 324). In keeping with this definition, he has identified three levels of visual virtualization that may be found in computer-generated media:

1)<u>Virtual Space:</u> When one or more pictorial depth cues are presented on a graphics display, the viewer perceives a three-dimensional spatial layout beyond the frame of the device.

2)<u>Virtual Image:</u> The observer can experience a sense of being incorporated into the virtual space when accommodative, binocular convergence and stereoscopic cues are provided in the display.

3)<u>Virtual Environment:</u> When fully implemented, this displayed information can evoke a sufficiently powerful sense of being in the virtual world that vestibular-ocular reflexes, vengeance and optokinetic reflexes consistent with depth relations in that world can occur. The frame of reference imposed by the medium is no longer salient to the observer.

The development of virtual environments for training applications is only in its infancy. As a result, there has been little research designed to evaluate the training effectiveness of VR environments in comparison to more conventional orthographic or perspective displays presented in a standard workstation with a computer monitor as the display device. The study described here was developed to systematically evaluate various display formats in a training situation. These displays include two-dimensional orthographic projections, three-dimensional perspective renderings of virtual space, and stereoscopic virtual environments. In this study, the training displays are informationally equivalent. That is, they all contain the same spatial information for learning to occur. They are not computationally equivalent. In turn, their computational non-equivalence affects the proximity compatibility of each display with the nature of the to-be-learned task. In some cases the orthographic displays are low in their compatibility with the to-be-learned task. In other cases, they are very compatible. The compatibility relationship is reversed for the perspective displays. Finally, among the perspective displays, their computational equivalence varies as a function of their immersive characteristics, from a conventional virtual space display on a standard computer monitor to a state-of-the art virtual environment display.

The Spatial Learning Problem

The task of interest in this training program was to determine the three dimensional spatial relationship between two aircraft from information displayed on the cockpit instruments of a fighter aircraft. In flying air-to-air intercepts, a fighter pilot must plan and execute most tactical maneuvers well before acquiring visual contact. Unquestionably, the rapid acquisition of an accurate mental model of the target's position is critical. To aid the pilot in building a mental model representing this three-dimensional information, the cockpit instruments present an array of spatial information. Figure 1 shows a stylized Head-Up Display (HUD) of a fighter aircraft. From the figure, it should be clear that this instrument is not readily interpretable by the casual observer. The spatial relationships represented in the data do not map onto a representation of three-dimensional airspace in any direct or intuitive manner. To build a mental model, relevant information must be discriminated in a cluttered display when it also is segregated spatially from correlated information about the same object. Three dimensional spatial relationships are represented alphanumerically, as well as in two-dimensional analog and digital formats and referenced to different coordinate frames. Because this information can be difficult to interpret and use, operators need to be trained to discover and implement strategies for creating a mental model of the situation.

For this study we developed a system for teaching students how to interpret and integrate spatial information presented on a simulated HUD display. Four properties of the momentary spatial relationship between two aircraft are relevant to specifying position. They are:

1) <u>range</u>: the horizontal distance separating the two aircraft;

2) <u>altitude</u>: the vertical distance separating the two aircraft;

3) <u>azimuth</u>: their angular separation in the horizontal plane;

4) <u>aspect:</u> their relative orientation.

The training objective of this study was to enable a student to translate HUD information specifying these properties into the actual location of a model aircraft in three-dimensional space. The HUD represented a target positioned somewhere in front of the student at a range of 10 miles, well beyond visual range. The spatial problem was to determine the expected location of the target. The experiment combined azimuth, altitude and aspect information to identify the target's position and orientation in space.

To aid in visualizing the target's relative spatial position, each student was supported by one of five visualization formats: two-dimensional orthographic projection(s), a three-dimensional perspective rendering, a wide Field-of-View (FOV) three-dimensional virtual environment and two narrow FOV virtual environment displays. Upon completing a training session under one of these visualization conditions, the student was tested on a spatial identification task. The research goal was to determine whether the different visualization aides would result in demonstrable differences in performance on the spatial problem-solving tasks. These displays may be compared on a variety of levels, but in this study three questions were of particular interest:

1. Virtual Environments versus Conventional Systems. Can students using state-of-the-art virtual environment training systems to visualize spatial relationships learn to solve spatial problems more effectively than subjects trained with conventional computer systems?

2. Display Parameters of the Virtual Worlds. Because virtual environment training systems are new, we do not know what minimal design parameters are necessary for assuring their functionality in training applications. Their success or failure in training may depend on specific characteristics of the display. Are there specific parameters of virtual environment displays systems, namely display opacity and field-of-view, that effect the utility of these media in training spatial problem solving skills? In this study, we sought to address these issues by creating three different virtual world training environments, a wide FOV opaque display and two narrow FOV displays, one opaque and the other transparent.

3. 2D versus 3D Visualization Formats. Finally, many conventional training systems offer only a two-dimensional, orthographic view of the to-be-visualized materials. However, as noted above, it is generally assumed that three-dimensional graphics are more effective for training three-dimensional spatial tasks, because of the compatibility of task and training format. Are three-dimensional formats more effective for training students to develop mental models of a three-dimensional spatial problem, or are two-dimensional formats informationally and computationally equivalent?

METHOD

The training process was conducted in three phases as described below.

Phase 1: Pre-Training Tutorial

Pre-training consisted of a short multimedia tutorial to introduce the subject to the HUD display, a description of the to-be-learned spatial properties and how to read and interpret the relevant spatial information on the HUD. Each tutorial was composed of a set of self-paced lessons in which text-based instruction was elaborated with graphics and video clips to emphasize and clarify critical spatial concepts.

Before completing each section of the tutorial, the subject was given a short multiple choice test set with feedback. The test items were designed to ensure that the subject could correctly identify the relevant data on the HUD display. Subjects were allowed to study the tutorial at their own pace, and to review as often as desired. Pre-practice training on the tutorial lasted about 45 minutes for each subject. Upon completing the tutorial, the subject was shown a sample HUD and learned how to control an object’s left/right movement (azimuth), up/down (altitude), and rotation (aspect) using the SpaceBall.

There was no final test for comprehension before continuing onto the second phase of training.

Phase 2: Spatial Localization Practice

Upon completion of the tutorial, subjects practiced mapping HUD symbology onto a representation of a target aircraft in a "pick and place" task. In this task, subjects were shown a sample configuration of spatial information on a HUD and a model of an aircraft. The subjects were told to imagine that they were pilots of an aircraft suspended in space at an altitude of 20,000 feet. The model aircraft was positioned straight ahead (azimuth = 0°), also at an altitude of 20,000 feet. The aspect angle of the target relative to the subject was 90L (left wing facing the subject). The HUD information represented a target at a range of approximately 10 miles, but the azimuth, altitude and/or aspect information did not correspond to the model's initial position. Although subjects/pilots normally would not be able to see a target at this distance, the task was to position the model at its expected location and/ or orientation to match the position described on the HUD. To position the target, the subject manipulated a SpaceBall™ a 3D input device. The 6 degrees-of-freedom of movement on the SpaceBall™ were constrained to allow the subject to move the target only in the spatial dimensions under investigation. Target placement was visually monitored by viewing the moveable model in one of the five visualization display formats being tested throughout the study. When satisfied with the position of the target, a keypress on the SpaceBall "froze" the model at the selected position, recorded this position and displayed a 'phantom' model at the target's correct spatial location for that HUD. Subjects were allowed to study and compare the positions of the moveable and feedback models for as long as desired. A second keypress removed the feedback 'phantom' and returned the moveable model to its initial position. The next HUD in the training set was then presented and subjects repeated the "pick and place" task for the next trial. Sets of training trials were blocked into training sessions and separated by short rest periods. Subjects were allowed as much time as they needed on each trial, but were instructed to work as fast as they could while trying to be as accurate as possible. The combination of 7 azimuth angles (0°, ±15°, ±30° and ±45°), 6 altitudes (0, 10,000, 20,000, 30,000, 40,000, and 50,000 feet) and 12 aspect angles (0, ±30°, ±60°, ±90°, ±120°, ±150° and 180°) produced a total of 504 unique combinations. This set was sampled randomly for 84 practice trials. The practice trials were grouped into 4 blocks of 21 trials and separated by a short rest period. Two measures of performance on each trial were recorded, response time and difference between the target's true and judged positions.

The "virtual world" in which the "pick and place" task was performed consisted of 40 square miles of flat, textured terrain. Other than the textured pattern of the terrain, there were no distinguishing landmarks furnishing the world. The designated target was a simple 3D model of an F-16, colored red. The 'phantom' feedback model was the same model, colored blue.

Phase 3: Transfer of Training Assessment

Upon completion of the training trials, the subject's skill at interpreting HUD symbology was assessed using a 2 alternative forced-choice recognition task. For this task the subject was again asked to imagine looking out the front window of an aircraft at 20,000 feet altitude. To simulate the out-the-window view in a manner not experienced in any of the practice conditions, the visual imagery was projected onto a 16 ft (h) x 6 ft (v) screen mounted on the wall 7 feet in front of the subject. This wide screen projection was created using two Sharpe XG-2000U™ Liquid Crystal Display (LCD) color projectors (640 x 480 resolution each projector). The projectors were positioned to fill the screen area and provide a 90° (h) x 30° (v) (approx.) FOV. A transparent HUD display was superimposed on this display using a 35mm slide. In this task, a 3D perspective image of a single target was projected on the wide-screen display along with the sample HUD. On half the trials, the target's spatial location accurately corresponded to the HUD information. In the remaining trials, the target's position did not correspond to the information displayed on the HUD. The subject's task was to judge (match/ no match) whether the target's location corresponded to the HUD information. Responses were recorded on the SpaceBall™ with one button indicating 'match' and another button indicating 'no match'. There was no feedback given during testing and the target did not return to the initial position. Each response triggered the next trial. Subjects were informed of the probability of a mismatch at the beginning of the test session. The recognition test set contained 48 HUDs and was constructed to allow eventual analysis of two variables. In this study, the test set contained 24 items from the practice set (OLD items) and 24 HUDs not seen during practice (NEW items). The second variable was the type of error introduced. Each of OLD/NEW item sets contained 12 test trials in which the target's position accurately depicted the position described by the HUD. Of the 12 mismatched trials, 4 contained an azimuth error, 4 contained an altitude error and 4 contained an aspect error. In the OLD item set, two of each error type involved a change in the HUD and two changed the target's position. In the final test, the subject was informed that some trials involved a mismatch, but no information regarding the nature of the mismatch was revealed. Again the subject was instructed to balance speed and accuracy in their responses. Performance was evaluated with two measures, response time and error rate. These final performance measures were intended to test the subject's ability to transfer skill at positioning a target to a more realistic spatial problem encountered during flight: based on HUD information, the pilot must predict the target's location and look for it at that predicted point in space; is it where you predict it to be?

Display Devices

The virtual world was generated by an XTAR Falcon_PC™ personal computer. This system was capable of producing a pair of stereoscopic images at a fill rate of 160 million pixels per channel per second. The XTAR system also was used to control the SpaceBall and to record data. In the experimental conditions using head-mounted displays, the XTAR also monitored head movements detected by an Ascension Flock of Birds™ (a magnetic sensing device) and updated the imagery to correspond to changes in head position. The imagery generated by the XTAR system was projected into one of five experimental display devices.

2D Orthographic Projections. The 2D orthographic condition (2D) (Figure 2a) consisted of a plan and elevation view of the virtual world projected onto a 13" Mitsubishi RGB color monitor (resolution 800 H and 560 V non-interlaced). Both the plan and elevation views were presented on the same monitor simultaneously. The screen display area was partitioned into a plan view that occupied a 100° (h) x 50° (v) FOV range and the elevation view occupied a 100° (h) x 25° (v) FOV. The model aircraft was depicted in both views, but to avoid confusion between the two displays, the models were presented as simple dots in the elevation view. This modification also eliminated any potential confusion due to the confounding introduction of the perspective into the display.

3D Perspective Projections. The 3D perspective condition (3D) (Figure 2b) consisted of a perspective view of the virtual world with the observer's eyepoint set at 20,000 feet altitude and directed straight ahead. Various cues, including interposition, relative size, relative brightness, height in plane also were incorporated in the display to provide depth and distance information. This eyepoint corresponded to the center of the 13" Mitsubishi RGB color monitor. The field of view represented in this projection measured 100° (h) x 75° (v) FOV. The initial position of the moveable model was at the center of this viewport and represented a target position of 00 azimuth, 20,000 feet altitude and 10 nautical miles range from the subject's aircraft. As with the orthographic display, the azimuth range of ±45° occupied nearly the entire horizontal viewing area. The vertical range of 0 - 50,000 feet occupied approximately two-thirds of the vertical display area.

Virtual Environment 1: Wide Field-of View Opaque Head-Mount. Three separate head-mounted displays were employed to explore two design characteristics of virtual environment training formats, field-of -view and display opacity. The first virtual environment condition, Wide FOV/Opaque Mount (VE1), represents a prototypical display format currently available commercially and offered on many turn-key systems (Figure 2c). In this condition, the imagery, in NTSC format, was projected into a Flight Helmet™, a stereoscopic head-mounted display composed of a pair of LCD displays (360 x 240 non-interlaced resolution) viewed through LEEP™ wide-angle optics. The combined images seen through these optics produce a FOV 90° (h) x 30° (v) (approx.). With this FOV, it was possible for the subject to observe the entire range of azimuth and altitude values used in these experiments without moving the head. However, the subject's unrestricted head movements were monitored by the magnetic head tracking system and the imagery was updated to correspond to the changing viewpoint in the virtual world. To further capitalize on the potential 'immersive' characteristics of the virtual environment display, the spatial layout of the virtual world was mapped onto the physical layout of the experimental setting. Thus, the subject could employ both visual and kinesthetic cues to judge a target's position. For example, suppose the HUD indicated that the target model was positioned at an azimuth angle of 45° right. A rightward head rotation of 45° would place the target aircraft straight ahead in the direction of gaze and centered in the display.

Virtual Environment 2: Narrow Field-of View Opaque Head-Mount. The second virtual environment condition, Narrow FOV/Opaque Mount (VE2), introduced a slight modification to the prototypical display format (Figure 2d). In this condition, the imagery again was projected into the Flight Helmet^{TM} , but it was magnified to effectively reduce the size of the field of view. Thus, the combined images seen through these optics produced a FOV 40° (h) x 14° (v) (approx.). With this considerably restricted FOV, the subject was required to move the head in order to observe the entire range of azimuth and altitude values used in these experiments. This introduced a peculiar problem for the subject when the azimuth or altitude value displayed on the HUD varied greatly from the initial position-remembering the location of the target when it was outside the field-of -view. For example, suppose the HUD azimuth is 45° right. Because the spatial layout of the virtual world was mapped onto the physical layout of the experimental setting, a rightward head rotation of 45° would center the target in the direction of gaze. Because the target's initial position was 0° azimuth, however, the head movement would have caused the target to slip out of the field-of-view. The subject must move the model into the center of the display based on its perceived distance from the visible edge of the display. This could be done easily if the gaze direction does not wander. However, if the gaze wanders and the subject does not attend to the change in head position, it would be very easy to become disoriented, or to lose the target in this relatively featureless virtual world. Experimentally, the inclusion of this condition was to study the effect of field of view on spatial problem solving in a virtual world training system.

Virtual Environment 3: Narrow Field-of View Transparent Head-Mount. The final virtual environment condition, Narrow FOV/Transparent Mount (VE3), introduced a third variation of the prototypical display format. In this condition, the imagery was projected into a high-resolution head-mounted display consisting of a pair of monochrome CRT displays (1280 x 1024 resolution) and associated optics. The combined stereoscopic images produce a 40° (h) x 30° (v) (approx.) FOV. As with the Narrow FOV/Opaque Mount (VE2), a subject wearing this display was required to move the head in order to observe the entire range of azimuth values used in these experiments. Head movements also would cause the target to slip into or out of the field of view. However, unlike the opaque condition, the transparent displays allowed the subject to use physical elements in the experimental setting to serve as cues, or anchors, for positioning the target. For example, with a 45° head rotation to the right, the straight ahead direction of gaze was directly at the corner of the room. The diagonal alignment of the floor tiles also corresponded to the 45° azimuth angle. Therefore, although the narrow field of view caused the target to slip out of visible range with ahead rotation, disorientation in the virtual world should be less troublesome if the subject aligned the target with ambient cues in the physical setting. Of interest in this condition was whether the availability of correlated cues in the physical and virtual worlds could facilitate three-dimensional spatial problem solving.

Subjects

Thirty-five subjects (20 male, 15 female), ranging from 17 to 45 years of age, participated in this study. No prior spatial training was found, and no subject had previous flying training. Seven subjects (4 male, 3 female) were randomly assigned to each of the five experimental conditions. Most subjects were trained individually rather than in pairs. All subjects were volunteers and were paid for their participation. As with the previous studies, no subject was familiar with virtual environment systems prior to training, but most reported previous experience with video games.

RESULTS

Performance during practice was assessed in terms of response time and judged azimuth, aspect and altitude of the target. For the Log2 transformation of each position measure, the absolute difference from the specified HUD value was computed as well as the signed differences. Separate analyses of variance were conducted on the transformed data for the azimuth, aspect and altitude differences. Also as with the previous studies, three planned contrasts were explored. These contrasts were: 1) comparison to the two monitor conditions versus the three head-mount conditions; 2) comparison to the 2D monitor condition versus 3D conditions; and, 3) comparison of the wide FOV head-mount condition versus the two narrow FOV conditions.

Practice Block Effects

Analysis of the absolute differences revealed significant effects of practice block for azimuth, aspect and altitude measures. Errors in positioning the target's azimuth [$F(3,30) = 11.64$, $p<0.0001$], aspect [$F(3,30) = 21.44$, $p<0.0001$] and altitude [$F(3,30) = 19.53$ $p<0.0001$] decreased for all conditions across the practice blocks. Of interest is the fact that for all three measures, performance appears to have leveled off at the third practice block Mean errors did not continue their downward trend in the fourth block of trials. The mean response time also decreased between the first and third practice blocks [F(3,30) = 94.07, p<0.0001], and leveled off thereafter. There was no significant interaction between block and condition. The learning curves were comparable among all five conditions.

Azimuth Errors

There was a main effect of condition in the analysis of variance [F(4,30) = 3.36, p<0.0218]. The mean errors in azimuth settings for the five conditions were $Mean_{2D}$ = 3.94 degrees, Mea_{n3D} = 5.24 degrees, $Mean_{VE1}$ = 5.74 degrees, $Mean_{VE2}$ = 6.87 degrees and $Mean_{VE3}$ = 5.28 degrees. The 3D monitor group errors remained lower than those of the virtual environment group and statistical differences in the planned contrast of the monitor versus the virtual environment conditions were significant. Subjects in the three head-mount conditions still produced greater error in setting the target's azimuth angle [F(1,30) = 7.03, p<0.0127]. The greatest difference between conditions, however, was revealed by the statistical contrast of the 2D monitor versus all other conditions [F(1 ,30) = 9.43, p<0.0045]. There were no differences in the planned contrast of the wide and narrow FOV head-mount conditions.

Subjects in the two narrow FOV head-mount conditions tended to overshoot the azimuth position while subjects in the other conditions had a slight tendency to undershoot. These differences were evidenced in a significant effect of condition [F(4,30) = 7.56, p<0.0002], as well as significant contrasts between the monitor versus head-mount conditions [F(1 ,30) = 15.07, p<0.0005], the 2D versus all others [F(1,30) = 9.51, p<0.0044] and, most importantly, the wide FOV versus the two narrow FOV head-mount conditions ($Mean_{VE1}$ = -1.28 degrees, $Mean_{VE2}$ = + 1.57 degrees and $Mean_{VE3}$ = +2.35 degrees). There was a slight condition effect [F(4,30) = 2.93, p<0.0369] which seems to be attributable to the difference between the 2D monitor and other conditions [F(1,30) = 4.61 , p<0.0401].

Aspect Errors

Subjects in the 2D condition produced more error in their aspect settings than subjects in the other display conditions ($Mean_{2D}$ = 21.56 degrees). The 3D monitor subjects were more accurate than the other groups ($Mean_{3D}$ = 11.24 degrees, $Mean_{VE1}$ = 13.36 degrees, $Mean_{VE2}$ = 15.24 degrees and $Mean_{VE3}$ = 14.83 degrees). Analyses of variance provided a statistically significant effect of condition [F(4,30) = 5.73, p<0.0015] as well as a significant difference in the planned contrast of 2D versus the other conditions [F(1 ,30) = 17.10, p<0.0003]. Subjects in the 2D condition also were less consistent in their settings. Analyses of the standard deviations of the signed aspect differences provided a statistically significant effect of condition [F(4,30) = 6.66, p<0.0006] as well as a significant difference in the planned contrast of 2D versus the other conditions [F(1,30) = 18.81, p<0.0001]. The measure of bias, mean signed aspect difference, indicated that subjects in the two narrow FOV head-mount conditions had a tendency to overshoot the aspect angle. The other three conditions showed no such tendency. The planned contrasts of monitor versus virtual environment displays evidenced this tendency [F(1 ,30) = 4.40, p<0.0445], as did the contrast of the wide versus narrow head-mount display conditions, which approached significance at F(1,30) = 3.58, 0<0.06 ($Mean_{VEI}$ = +1.24 degrees, $Mean_{VE2}$ = +1.89 degrees and $Mean_{VE3}$ = +2.04 degrees).

Altitude Errors

Analysis of the absolute altitude differences failed to achieve a statistically significant effect of condition or in any of the planned contrasts. Visual inspection of the data, however, indicated that the virtual environment display conditions produced slightly greater absolute error than the monitor conditions ($Mean_{2D}$ = 3,864 feet, $Mean_{3D}$ = 3, 732 feet, $Mean_{VE1}$ = 4,500 feet, $Mean_{VE2}$ = 3, 972 feet and $Mean_{VE3}$ = 3, 655 feet). The measure of bias, mean signed altitude difference provided significant differences among conditions [F(4,30) = 3.49, p<0.0186] and between the wide versus narrow FOV conditions [F(1,30) = 11.11, p<0.0023]. The bias measure indicated a tendency to overshoot altitude in the monitor conditions and in the wide FOV head-mount display condition ($Mean_{VE1}$ = +1 ,592 feet); there was a slight bias to undershoot in the narrow FOV virtual environment conditions ($Mean_{VE2}$ = -1,149 feet and $Mean_{VE3}$ = -1,378 feet). Finally, the measure of subject consistency, standard deviation of the signed altitude difference, indicated no differences among conditions.

Practice Response Time

The analysis of variance of mean log response times revealed that subjects in the different conditions required differing amounts of time to complete the pick and place task [F(4,30) = 8.83, p<0.0001]. Response times were fastest for subjects in the 3D monitor condition ($MeanRT_{3D}$-Monitor = 19.97 s) and slowest for the opaque narrow FOV head-mount display condition ($MeanRT_{VE2}$ = 40.22 s). The other conditions were distributed in this range in a fairly systematic fashion ($MeanRT_{2D}$-MonItor = 23.26 s, $MeanRT_{VE1}$ = 24.95 s and $MeanRT_{VE3}$ = 36.76 s). A significant effect in the planned contrast of monitor versus virtual environment conditions [F(1,30) = 21.21, p<0.0001] also demonstrates that the monitor conditions allowed the fastest response times. The clearest distinction among conditions, however, was evident in the planned contrast of the wide FOV versus narrow FOV head-mount display conditions [$F(1,30) = 12.11$, $p<0.0016$]. Clearly, subjects in the narrow FOV conditions required more time to complete the pick and place task.

Test Performance

Test performance at correctly discriminating the matched/mismatched targets and HUDs were conducted. Subjects in the wide FOV head-mount display condition attained the highest accuracy, but the differences among practice conditions did not achieve statistical significance. The planned comparison of wide FOV versus narrow FOV displays approached significance [$X^2(1,30) = 2.75$, $p<0.09$]. The contrast of 2D monitor versus other 3D conditions produced a significant difference [$X^2(1,30) = 3.898$, $p<0.O48$], indicating that subjects in the 2D monitor condition performed significantly less well than subjects in the 3D practice environments. There were no differences among conditions in response time to the test items ($MeanRT_{2D} = 6.59$ s, $MeanRT_{3D} = 6.11$ s and $MeanRT_{VE1} = 6.77$ S, $MeanRT_{VE2} = 6.58$ s and $MeanRT_{VE3} = 7.11$ s).

Performance accuracy varied considerably as a function of the type of error presented in the trial. Subjects were most accurate at recognizing target/HUD matches. Mismatches between the HUD and model's azimuth angle or altitude produced comparable accuracy. Performance on trials containing mismatches between the HUD and model aspect angles again was reduced to chance. This dramatic drop in performance when aspect angle errors were introduced was statistically significant [$X^2(3,30) = 91.25$, $p<0.000$. Responses to mismatched aspect angles took slightly longer than responses to the other matched and mismatched pairs [$F(3,30) = 8.43$, $p<0.0001$].

A final planned comparison explored the effect of introducing new target/HUD pairs versus pairs that had been encountered during practice. Performance on never-seen-before items was clearly better than on items with which the subjects had prior experience [$X^2(1,30) = 21.033$, $p<0.000$]. There were no differences in response times to the new and old items, however.

DISCUSSION

In the introduction, we discussed some taxonomies of visualization media which may serve to identify those displays with the greatest potential for learning to visualize three dimensional spatial relationships. Wickens and colleagues have proposed that media be classified in terms of the spatial coincidence, or correspondence between the format of the display and task requirements. According to this framework, those displays with the greatest dimensional similarity to task demands stand the highest probability of successfully serving the student. Gibson has pointed out that non-representational information (such as the display housing) can afford the student with greater or lesser access to the to-be-learned information . Because the medium itself is inescapably a part of the training environment, Gibson and, more recently, Ellis have suggested a taxonomy which is characterized by a continuum of 'virtualization.' For practical purposes, it is bounded by highly salient physical characteristics of the training medium itself (e.g., a set of two-dimensional markings framed by the display device) and by highly salient properties of the three-dimensional virtual world in which the student experiences psychological ‘presence.'

In this study, we have sought to explore a variety of these visualization media within a specific task environment. The goal of this research has been to examine the effectiveness of these media in learning a spatial positioning task; and in transferring this new knowledge to a recognition task. Throughout this study, we have sought to gain an understanding of three particularly engaging questions derived from the Wicken's and Gibsonian taxonomies. They are:

1) Are three-dimensional formats more effective for training students to develop mental models of a three-dimensional spatial problem, or are two-dimensional formats informationally and computationally equivalent?

2) Can students using state-of-the-art virtual environment training systems to visualize spatial relationships learn to solve spatial problems more effectively than subjects trained with conventional computer systems?

3) Are there specific parameters of virtual environment displays systems, namely display opacity and field-of-view, that effect the utility of these media in training spatial problem solving skills?

The spatial prob1em-solving skill of interest in this research program was the ability to judge a target's location in space relative to oneself . In light of the results of this study, let us address each of research questions in turn.

Are informationally equivalent two and three-dimensional display formats also computationally equivalent for learning to read and interpret spatial displays? Conventional wisdom has held that, when available training systems offering three-dimensional computer graphics would be more effective than orthographic displays for training three-dimensional spatial tasks. Task and training format would be dimensionally compatible, a decidedly desirable solution. The results of the present experiment support this widely held belief - to a certain extent. For setting the azimuth and altitude of the target, the 2D displays had the advantage because they offered a highly salient two-dimensional frame of reference. The subjects' positioning accuracy and response times indicated that they benefited from this advantage. This study required the 2D subjects to monitor a pair of orthographic displays to accurately position the target. Again, the 2D subjects seemed to adapt well to the procedure. The advantage offered by the 2D displays was diminished when setting the aspect angle. In this case, the frame of reference was no longer useful. Of particular interest was the finding that the angular resolution on the display monitor was too low to accurately discriminate the aspect angle rotation. Recall that these are 30° shifts - large enough to have serious consequences if misjudged by a fighter pilot.

The advantage found for the 2D subjects in the second phase of training also was undermined when these subjects were placed in the position of recognizing the target's position in a more ecologically valid test environment. Taken together, these results would suggest an interesting dilemma for designers of spatial problem-solving training systems. The use of 2D orthographic displays for training may lead to apparent success at locating the target in a positioning task. Unfortunately, this success may not be indicative of eventual success in the field.

The second question motivating this research dealt with the training efficacy of virtual environment display systems. Recent developments in the advanced technologies have fostered a growing interest in virtual environments as training tools for visualizing complex spatio-temporal relationships. While synthetic environments were formerly the province of an elite class of dedicated hardware, technological breakthroughs in the last few years have changed the medium dramatically. It is now possible to produce real-time interactive “Virtual environments” on a desktop workstation. The increased availability of these advanced technologies has generated sufficient excitement for some to predict that virtual reality represents a new era of visualization - one in which the training system purposively relieves the cognitive workload on the user by presenting critical information in a natural and intuitive interface. Although the same information could be provided under traditional orthographic views or perspective views incorporating pictorial depth cues, many tend to assume that virtual environments will have a significant impact on our ability to encode, retain and recall information. In the present research, it was intuitively appealing to hypothesize that the virtual environments were comparable to the desired outcome of training - a mental model of three-dimensional spatial relationships. Therefore, subjects in these conditions should have enjoyed an overall training advantage.

The results of the present research have indicated that students can learn to solve spatial problems effectively using state-of-the-art virtual environment training systems. Whether these systems are more efficient than conventional systems remains a matter of opinion . On the one hand, subjects practicing the pick and place task using the head mounted displays required more time to accomplish the task and produced greater error in the azimuth and altitude settings. In setting the aspect angles, these subjects did not suffer the distinct dis-advantage some of the monitor subjects experienced. Thus, it could be argued that overall consistency in solving spatial problems in all three dimensions was facilitated in the virtual environment conditions. This is no small accomplishment. We can only speculate that if allowed more practice trials, subjects in these conditions could also have decreased their errors and response times. Recall that for all subjects, this was their first experience with a head mount display. All had reported some experience with more conventional displays. Thus, at the outset there was a general knowledge base for interacting with the 2D and 3D formats. As virtual environment display systems become more commonplace, this gap in prior experience will diminish.

Poorer performance in the practice phase of training should indicate that subjects in the virtual environment conditions did not learn to read and interpret target position as welt as those in the conventional monitor conditions. However, the recognition test results do not support this conclusion. Overall, subjects in the virtual environment conditions performed as well or better than their counterparts using the conventional systems for practice. Although the performance differences were not dramatic, they would not have been predicted from the practice data. Clearly, despite difficulties during practice, these subjects were learning to solve the spatial problems. Again, one can only speculate that as the systems become more commonplace, training benefits also may become more evident. The results of the recognition task are particularly prophetic of future potential for virtual environment training systems. Subjects in the wide FOV/opaque condition performed the pick and place task with average skill. Nevertheless, their recognition performance in the transfer-of-training task was higher than all other conditions. These results would suggest that they had developed a level of understanding superior to that of subjects in the other conditions. Future research might reveal even greater potential for virtual environment training systems in the realm of complex three-dimensional spatial problem solving.

The final question addressed in this study explored specific parameters of the virtual environment training systems, display opacity and field-of-view. Display opacity offered subjects different cues for locating the target. In the opaque display conditions, subjects were required to rely on the meager cues provided by the terrain texture. Subjects wearing the transparent displays, could opt for the terrain features or physical cues in the experimental setting. The results have not revealed any apparent advantage or disadvantage of display opacity. Anecdotal evidence acquired in observing subjects using the head-mounted display conditions revealed some Interesting behaviors, however. In all of the virtual environment conditions, subjects appeared to ignore kinesthetic information while wearing the head mounted display. For example, the widest azimuth angle setting was ±45°. T o set the target at this position, one need only rotate the head 45° and position the model in the center of the display. Subjects commonly rotated their heads well past 45°; often as much as 90°. They did not seem to appreciate that the head rotation was excessive. Feedback from the phantom model failed to extinguish this behavior. Throughout the practice blocks, some subjects continued to ignore kinesthetic information.

Subjects in the transparent head-mounted condition also had salient cues in the physical setting with which to calibrate their target's position. Nevertheless, it was evident to the experimenter that they were not aware of the utility of these cues. Their behavior did not reflect the use of room cues to anchor spatial relationships in the virtual world to corresponding relationships in the physical world. Perhaps if they had been instructed on the utility of these cues, they could have been used to their advantage. However, this type of instruction rarely would be afforded to students in the training environment. If it is not a strategy that would emerge naturally in the course of training for most subjects, it probably has little training value.

The field-of-view provided in the head-mounted displays also appeared to have marginal impact on their training effectiveness. Observational evidence indicated that subjects in both narrow FOV conditions clearly had a more difficult time adapting to the training system. They also seemed more prone to systematic bias to over or undershoot the target's position. These biases would suggest that with the diminished FOV, they had less of an understanding of the three-dimensional space of the virtual world. Previous work by Beer (1993) also has indicated that subject's distort spatial relationships beyond the instantaneous FOV. The present results have failed to provide a clear picture of the nature of this distortion. Future research should be conducted to explore this problem with greater precision.

SUMMARY AND CONCLUSION

In summary, the results of this experiment have revealed interesting characteristics of learning spatial skills using different training media. It has become amply clear that a salient frame of reference offered by the display frame may be highly beneficial when learning certain spatial relationships. On the other hand, an integrated display format dimensionally compatible with the task also is highly beneficial. Furthermore, it would appear that a wide FOV virtual environment display may have distinct advantages for generalizing training - at least for complex spatial problem solving tasks. It could be that the inclusion of more salient cues for spatial reference marking could reveal substantial benefits for the virtual environment conditions. The results of the present research indicate that virtual environment visualization technologies can and should play an important role in the future of spatial problem-solving training systems. Certain problems remain, however. The display technologies are primitive compared to conventional systems in terms of display resolution, comfort and ease of use. Students are not as familiar or comfortable with the technologies. As a result they are restrained in their willingness to engage and exploit the systems (i.e., they appear to be less willing to try new strategies). As these problems are resolved, we should see further application of virtual environment technologies to training problems.

REFERENCES

Beer, J. (1993). Perceiving scene layout through an aperture during visually simulates self motion. Journal of Experimental Psychology: Human perception and P 19, 333-365.

Bertoline, G. R. (1991). Using 3D Geometric models to teach spatial geometry concepts. Engineering Design Graphics Journal,-55, 37 -47.

Brody, A. R., Jacoby, R., & Ellis, S. R. (1991, October). Man overboard! What next? Paper presented at the 42nd International Astronautical Congress, Montreal, Canada.

Cole, R. E., & Parker, D. L. (1989). Stereo TV improves manipulator performance. SE!E Vol. 1083: Three-Dimensional Visualization and Display Technologies. (pp. 19-27). Bellingham, WA: SPIE.

Cutting, J.J (1988). Affine distortions of pictorial space: Some predictions for Goldstein (1987) that La Gournerie (1859) might have made. Journal of Experimental psychology: Human perception and Performance, 14, 305-311.

Cooper, L. A., Mowafy, L., & Stevens, J. J. (1986, November). Perceiving hidden structure: Viewindependent recognition of constructed three-dimensional representations. Paper presented at the 27th annual meeting of the Psychonomic Society , New Orleans, LA.

Ellis, S. R. (1991). Nature and origins of virtual environments: A bibliographical essay. Computing Systems in Engineering, .2, 321-347.

Farrell, E. J., & Christidls, Z. D. (1989). Visualization of complex data. SPIE Vol. 1083: three-dimensional Visualization and Display Technologies. (pp. 153-160). Bellingham, WA: SPIE.

Fisher, S. S., & Tazelaar, J. M. (1990). Living in a virtual world. ant, 15,215-221

Foley, J. D. (1987). Interfaces for advanced computing. Scientific American, ~, 127-135.

Gibson, J. J. (1979). The ecological acgroach to visual cercemion Boston, MA: Houghton Mifflin.

Goldstein, A. (1987). Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. Journal of Experimental Psychology: Human Perception and Performance, 13., 256-266.

Goldstein, A. (1988). Geometry or not geometry? Perceived orientation and spatial layout in pictures viewed at an angle. Journal of Exprimental Psychology: Human Perception and Performance, 14, 312-314.

Gomez, J. E. (1989). Scientific work environments in the next decade, SPIE Voi. 1083: Three-Dimensional Visualization and Disclax Technologies, (pp. 234-239). Bellingham, WA: SPIE.

Greenberg, D. P. (1991). Computers and architecture, Scientific American. 2.4Z, 104-109.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. Cognitive Science, 11, 65-99.

Liu, Y., & Wlckens, C. D. (1992). Use of computer graphics and cluster analysis in aiding relational judgment. Human Factors, M, 165-178.

McConathy, D. A., Deirdra, M. & Doyle, M. (1991). Interactive displays in medical art. In S. R. Ellis, M. K. Kaiser, & A. C. Grunwald (Eds.) Pictorial Communication in Virtual and Real Environments. (pp.97-110). New York: Taylor Francis.

Russell, G., & Miles, R. (1989). Volumetric visualization of 3D Data. In Ellis, S. R. , Kaiser, M. K. and A. Grunwald, (Eds.). Spatial Di§glaxs and Spatial lnstruments, (pp. 48/1 48/7). NASA Conference Publication #10032.

Wickens, C.D., & Andre, T. (1988). Proximity compatibility and the object display. Proceedings of the Human Factors SocietY 32th Annual Meetina (p. 1335-1339).

Wickens, C.D., & Todd, S. (1990). Three Dimensional display technology for aerospace and visualization. Proceedings of the Human Factors Society 34th Annual Meeting. (p. 1479-1483).

Wickens, C. D., Todd, S., & Seidler, K. (1989). Three-dimensional displays: Perception. Implementation. and Acollcation. (CSERIAC SOAR-89-01). Wright-Patterson AFB, OH: Armstrong Aerospace Medical Research Laboratory.

Zeltzer, D. (1992). Autonomy, interaction and presence. Presence,1, 127-132.

Zsombor-Murray, P. J. (1990). 2-D and 3-D CAD: Complements to visualization. Engineering Design Graphics Journal. M, 17-29.